UDC 551.509.617:551.501.81:551.577.11:519.241.2

USE OF THE GAMMA DISTRIBUTION IN SINGLE-CLOUD RAINFALL ANALYSIS

JOANNE SIMPSON—*Experimental Meteorology Laboratory, Environmental Research Laboratories, NOAA, Miami, Fla.*

ABSTRACT—This study is based on the radar-evaluated rainfall data from 52 south Florida cumulus clouds, 26 seeded and 26 control clouds, selected by a randomization procedure. The fourth root of the rainfall for both seeded and control populations was well fitted by a gamma distribution for probability density. The gamma distribution is prescribed by two parameters, one for scale and one for shape. Since the coefficient of variation of seeded and control cloud populations was the same, the shape parameters for the two gamma distributions were the same, while the seeded population's scale parameter was such as to shift the distribution to higher rainfall values than the control distribution. The best-fit gamma functions were found by application of the principle of maximum entropy.

Specification of tractable distributions for natural and modified rainfall populations provides an important prerequisite for the evaluation of seeding effects by Bayesian statistics, a continuing objective in the Experimental Meteorology Laboratory cumulus seeding programs.

In meteorology, the gamma function has been used extensively to fit rainfall data on fairly large space and time scales, ranging from individual storms up to monthly and yearly distributions (Thom 1958, 1968, and references therein). Recently, the Experimental Meteorology Laboratory (EML) has found this function of great value in analyzing radar-evaluated rainfalls from single Florida cumulus clouds, both natural and seeded.

In 1968 and 1970, EML conducted randomized pyrotechnic seeding experiments on single clouds in south Florida. The seeding was done massively, to release the heat of fusion latent in the supercooled water and, under predictable conditions, to cause the seeded clouds to grow larger and process more water than their unseeded counterparts. A one-dimensional model (Simpson and Wiggert 1971) was run in real time to predict seedability(excess vertical growth of seeded clouds). The rainfall analyses were made using the University of Miami 10-cm calibrated radar (Woodley 1970) and results were tested against a rain gage network (Woodley and Herndon 1970). Altogether, 26 seeded and 26 control clouds were compared (Simpson et al. 1971). Statistical analyses showed that the seeding effect on rainfall exceeded a factor of three; there was a mean seeded minus control difference of about 270 acre-feet per cloud. The statistical significance of the difference was better than 0.05.

In connection with a Bayesian analysis of more complicated sequel experiments (Simpson and Pézier 1971), it became necessary to seek a tractable distribution function; that is, a function with finite, readily specified moments to fit these single-cloud rainfall data. While this effort has so far encountered obstacles when the raw data are used, the gamma function was found extremely useful in treating their fourth root, or the "transformed" data, as we shall show briefly here The idea of transforming these data by taking their fourth root was introduced by G. F. Cotton (see appendix, Woodley et al. 1970) in his statistical analyses, for the purpose of making certain significance tests more applicable. The raw and transformed data are presented in table 1; their origin and limitations have been described elsewhere (Woodley et al. 1970).

The gamma probability density function may be written

$$p(R)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}R^{\alpha-1}e^{-\beta R} \tag{1}$$

where $p(R)$ is the probability density of a rainfall amount, R (here measured in acre-feet), from a single cloud. The scale of the distribution is determined by the parameter β and the shape by the parameter α. Γ is the gamma function (Pearson 1951). The first two moments of the gamma function are well known to be (e.g., Tribus and Pézier 1970, Kendall and Stuart 1963)

$$\mu_1=\langle R\rangle=\alpha/\beta \tag{2}$$

and

$$\mu_2=\sigma^2=\alpha/\beta^2 \tag{3}$$

where $\langle R\rangle$ is the expected value and σ^2 is the variance. Therefore, the coefficient of variation, V, is

$$V\equiv\frac{\sigma}{\langle R\rangle}=\frac{1}{\sqrt{\alpha}} \tag{4}$$

The best-fit gamma functions to the transformed data in table 1 are found by application of the principle of maximum entropy. Tribus (1969, p. 197) has shown that this method, readily computerized, gives better fits than the classical chi-square approach, which is in fact an approximation to the maximum-entropy method. The same principle is applied to find the best fit to the data of six other well-known functions, as listed in table 2, and then the excellence of the fits are compared by means of

TABLE 1.—*Single-cloud data for 1968 and 1970*

Total cloud lifetime: Seeded rain, Acre-feet	Seeded rain, Fourth root	Control rain, Acre-feet	Control rain, Fourth root
129. 6	3. 37405	26. 1	2. 26027
31. 4	2. 36719	26. 3	2. 26459
2, 745. 6	7. 23868	87. 0	3. 05408
489. 1	4. 70272	95. 0	3. 12199
430. 0	4. 55373	372. 4	4. 39291
302. 8	4. 17147	0	0 (1)
119. 0	3. 30283	17. 3	2. 03944
4. 1	1. 42297	24. 4	2. 22253
92. 4	3. 1004	11. 5	1. 84151
17. 5	2. 04531	321. 2	4. 23344
200. 7	3. 76389	68. 5	2. 87689
274. 7	4. 07113	81. 2	3. 00185
274. 7	4. 07113	47. 3	2. 6225
7. 7	1. 6658	28. 6	2. 31255
1, 656. 0	6. 37918	830. 1	5. 36763
978. 0	5. 59223	345. 5	4. 31134
198. 6	3. 754	1, 202. 6	5. 88885
703. 4	5. 14992	36. 6	2. 45963
1, 697. 8	6. 41906	4. 9	1. 48782
334. 1	4. 27532	4. 9	1. 48782
118. 3	3. 29797	41. 1	2. 53198
255. 0	3. 99609	29. 0	2. 3206
115. 3	3. 27686	163. 0	3. 57311
242. 5	3. 94619	244. 3	3. 95349
32. 7	2. 39132	147. 8	3. 48673
40. 6	2. 52424	21. 7	2. 15832

TABLE 2.—*Probability distribution functions*

Distribution	*Equation for* $p(R)$
Truncated normal	$A \exp(Bx - Cx^2)$
Gamma	$A\, x^B \exp(-Cx)$
Weibull	$A\, x^B \exp(-Cx^{B+1})$
Log-normal	$A\, x^B \exp(-C \log x^2)$
Rayleigh	$A\, x^B \exp(-Cx^2)$
Inverted gamma	$A\, x^B \exp(-C/x)$
Inverted Rayleigh	$A\, x^B \exp(-C/x^2)$

TABLE 3.—*Results of program* DAMAXS2 *for single-cloud transformed rainfalls*

I. Unseeded Cases

Distribution	Parameters: $-$Log A	B	C	Relative probability
Tr. normal	4. 38621	2. 26211	0. 383128	0. 0515
Gamma	0. 491139	5. 523	2. 2236	0. 271176
Weibull	2. 34166	1. 801	3. 43$\times 10^{-2}$	0. 0943
Log-normal	3. 08051	5. 15809	3. 08645	0. 256951
Rayleigh	2. 31008	2. 5878	0. 180015	0. 169842
Inverted gamma	$-$11. 7724	$-$7. 16194	15. 3737	0. 124435
Inverted Rayleigh	$-$4. 22864	$-$4. 21716	8. 40303	0. 0316

Distribution	χ^2	
Tr. normal	12. 78	16 Degrees of Freedom
Gamma	9. 46	
Weibull	11. 57	
Log-normal	9. 57	
Rayleigh	10. 39	
Inverted gamma	11. 02	
Inverted Rayleigh	13. 75	

II. Seeded Cases

Distribution	Parameters: $-$Log A	B	C	Relative probability
Tr. normal	5. 43878	2. 14824	0. 273406	0. 121835
Gamma	2. 47642	6. 10433	1. 83149	0. 244059
Weibull	3. 34546	2. 017	1. 168$\times 10^{-2}$	0. 183406
Log-normal	5. 34034	7. 34813	3. 25198	0. 142415
Rayleigh	3. 59768	3. 00345	0. 117279	0. 251999
Inverted gamma	$-$13. 8461	$-$7. 29359	20. 9373	0. 0467
Inverted Rayleigh	$-$ 5. 17874	$-$4. 23249	14. 9841	0. 0958

Distribution	χ^2	
Tr. normal	13. 02	16 Degrees of Freedom
Gamma	11. 63	
Weibull	12. 21	
Log-normal	12. 71	
Rayleigh	11. 57	
Inverted gamma	14. 94	
Inverted Rayleigh	18. 11	

the chi-square test and by relative probabilities. The latter are found using Bayes' equation, following the procedure described by Tribus (1969, p. 156 and ff.). All these calculations are made by means of a computer program entitled DAMAXS2, described and listed elsewhere (Simpson and Pézier 1971). The results are shown in table 3.

The first important result is that the gamma distribution is either best or a close second from either criterion, for both seeded and control data. The second important result is that α and V are nearly identical for both populations, namely

Unseeded

$$\alpha=6.16 \qquad V=0.40 \tag{5}$$
$$\beta=2.22$$
$$\langle R\rangle=\alpha/\beta=2.78$$

and

Seeded

$$\alpha=7.10 \qquad V=0.38 \tag{6}$$
$$\beta=1.83$$
$$\langle R\rangle=\alpha/\beta=3.88.$$

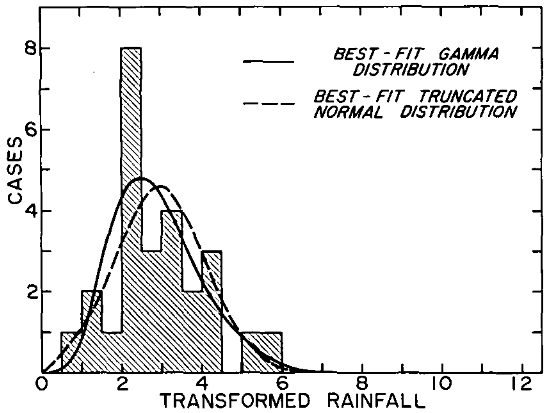

FIGURE 1.—Histogram and best-fit curves for the transformed rainfall (acre-feet) from the unseeded clouds. The grouping of the data is by intervals of 0.5. In the DAMAXS2 program, the group interval is $\sigma/3$.

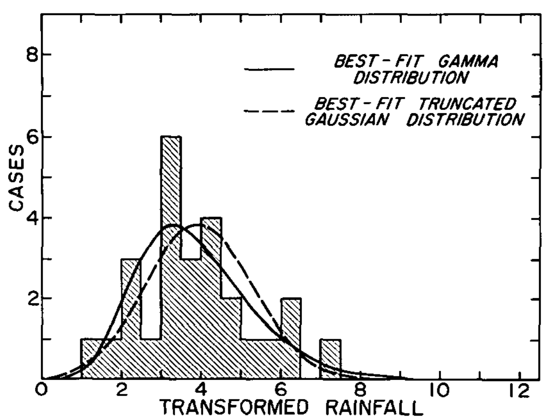

FIGURE 2.—Histogram and best-fit curves for the transformed rainfall (acre-feet) from the seeded clouds.

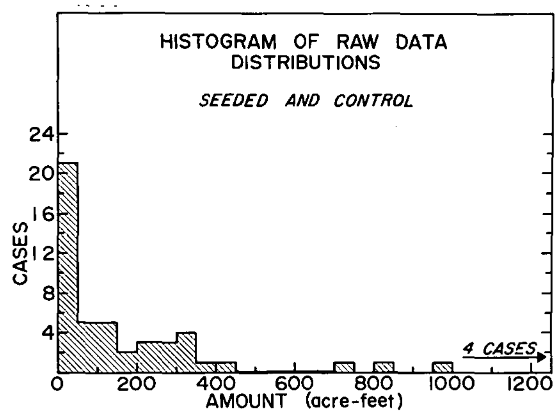

FIGURE 3.—Histogram of raw data distributions for 1968 and 1970 single clouds combined, 50-acre-feet subdivisions.

This result says that the seeding does not appreciably affect the shape or coefficient of variation [1] of the distribution but merely advances the mean of the transformed data by a factor (in this case about 1.4). In any meteorological experiment, if one can demonstrate or assume that the modification has this simple type of effect upon the data distribution (or on one of its roots), Bayesian analysis provides an easy and powerful tool for calculating the magnitude and significance of the modification, as we have shown elsewhere (Simpson and Pézier 1971).

Figures 1 and 2 show the data histograms with the best-fit gamma and truncated normal distributions superposed. Although table 3 shows that the truncated normal distribution has a χ^2 much smaller than the number of degrees of freedom for both data populations, it does not show well either the skewness or the mode of the data. In the case of the truncated normal distribution, a complicated relation exists between the parameters B and C in table 2 and the first two moments of the distribution. Tribus (1969, p. 171) provides a computer program [2] for finding these moments, given B and C. Using it, we found again that the best-fit curves for the seeded and unseeded cloud populations had very similar coefficients of variation; that is to say, the seeded population had a higher mean and a correspondingly higher standard deviation than the unseeded population. Clearly, the gamma distribution is most ideally suited to the analysis because the observed conservative property (V) of the unseeded and seeded populations is directly and simply related to the shape parameter, α, in the equation for the probability density.

Finally, a histogram of the raw data for seeded and control cases combined is shown in figure 3. The only distribution giving a good fit to these data is the inverted Rayleigh, which does not have finite moments and is not, therefore, very useful. Alternatively, we might usefully regard this graph as showing two separate populations, one with many members and small rainfalls, the other with few members and very large rainfalls. Consequently, an attempt is being made to treat each raw data set (seeded and control) as the sum of two gamma distributions.

ACKNOWLEDGMENT

The writer is grateful to Jacques Pézier for the DAMAXS2 program, which made this study possible.

REFERENCES

Kendall, Maurice G., and Stuart, Alan, *The Advanced Theory of Statistics* (2d Ed.), Vol. 1, Distribution Theory, The Hafner Publishing Co., New York, N.Y., **1963**, 431 pp.

Pearson, Karl (Editor), *Tables of the Incomplete Γ Function*, Cambridge University Press, Cambridge, England, **1951**, 164 pp.

Simpson, Joanne, and Pézier, Jacques, "Outline of a Bayesian Approach to the EML Multiple Cloud Seeding Experiments," *NOAA Technical Memorandum* ERLTM-OD 8, Environmental Research Laboratories, Office of the Director, Boulder, Colo., June **1971**, 59 pp.

Simpson, Joanne, and Wiggert, Victor, "1968 Florida Cumulus Seeding Experiment: Numerical Model Results," *Monthly Weather Review*, Vol. 99, No. 2, Feb. **1971**, pp. 87–118.

[1] A direct calculation of V from the fourth root data of table 1 gives 0.43 and 0.37 for unseeded and seeded populations, respectively, or a difference of 13 percent. Using the raw data, we get 1.69 and 1.47, again a difference of 13 percent.

[2] A minor modification of print instructions in DAMAXS2 yields the results of the same calculation.

Simpson, Joanne, Woodley, William L., Miller, Alan H., and Cotton, Gerald F., "Precipitation Results of Two Randomized Pyrotechnic Cumulus Seeding Experiments," *Journal of Applied Meteorology*, Vol. 10, No. 3, June **1971**, pp. 526–544.

Thom, Herbert C. S., "A Note on the Gamma Distribution," *Monthly Weather Review*, Vol. 86, No. 4, Apr. **1958**, pp. 117–122.

Thom, Herbert C. S., "Approximate Convolution of the Gamma and Mixed Gamma Distributions," *Monthly Weather Review*, Vol. 96, No. 12, Dec. **1968**, pp. 883–886.

Tribus, Myron, *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York, N.Y., **1969**, 478 pp.

Tribus, Myron, and Pézier, Jacques, "Concerning the Economic Value of Experimentation in the Design of Desalting Plants," *Desalination*, Vol. 8, Elsevier Publishing Co., Amsterdam, The Netherlands, **1970**, pp. 311–349.

Woodley, William L., "Precipitation Results From a Pyrotechnic Cumulus Seeding Experiment," *Journal of Applied Meteorology*, Vol. 9, No. 2, Apr. **1970**, pp. 242–257.

Woodley, William L., and Herndon, Alan, "A Raingage Evaluation of the Miami Reflectivity-Rainfall Rate Relation," *Journal of Applied Meteorology*, Vol. 9, No. 2, Apr. **1970**, pp. 258–264.

Woodley, William L., Simpson, Joanne, Miller, Alan H., MacKay, Steven, Williamson, Richard, and Cotton, Gerald F., "Some Results of Single Cloud Pyrotechnic Seeding in Florida, 1970," *NOAA Technical Memorandum* ERLTM–AOML 10, Atlantic Oceanographic and Meteorological Laboratories, Miami, Fla., Nov. **1970**, 87 pp.

[*Received June 29, 1971; revised August 23, 1971*]